T
Univ
G

Bulletin of The University of Georgia

VOLUME XXXIII AUGUST, 1933 NUMBER 11c

PHELPS-STOKES FELLOWSHIP STUDIES

NUMBER 12

A COMPARISON OF THE INTELLIGENCE AND ACHIEVEMENT OF TWELVE-YEAR OLD NEGRO CHILDREN IN THE RURAL SCHOOLS OF CLARKE COUNTY, GEORGIA

LIBRARY
NOV 3 1933
UNIVERSITY OF TORONTO

By Sarah Thurmond, M. A.
Phelps-Stokes Fellow in The University of Georgia, 1931-1932

Price 50 cents

Entered at the Post Office at Athens, Ga., as Second Class Matter, August 31, 1905, under Act of Congress of July 16, 1904. Issued Monthly by the University

SERIAL NO. 544

THE UNIVERSITY OF GEORGIA PUBLICATIONS

The University of Georgia publishes the special series listed below:

ADDRESSES.

Delivered at commencements, or on the occasion of anniversaries or other special convocations.

BUREAU OF BUSINESS RESEARCH STUDIES.

Published on a fund provided by friends of the University.

GEORGIA PRESS INSTITUTE ADDRESSES AND LECTURES.

The University has printed only those delivered in 1929 on its campus (Bulletin No. 462, 1929, price 50 cents).

INSTITUTE OF PUBLIC AFFAIRS PROCEEDINGS.

A selection of the addresses delivered at annual sessions of the Institute. These deal with state, national and international problems and affairs. Addresses delivered at the first two sessions have not as yet been published, except addresses delivered at the second session by Sam. J. Slate (Bulletin No. 447, 1928, price 25 cents).

MARTIN REYNOLDS SMITH MEMORIAL PRIZE STUDIES.

Published results of research work in chemistry, for which this prize is awarded. Volume I, number 1, only published thus far. Further issues not planned at present.

PHELPS-STOKES FELLOWSHIP STUDIES.

The results of work done toward a "scientific study of the Negro and his adjustment to American civilization" by holders of the Fellowship established in 1912 by Miss Caroline Phelps-Stokes. *A list of those thus far published is given on the inside back cover of this study.*

STUDIES.

Original work done by members of The University of Georgia Faculty, thus far chiefly in the field of Southern history.

CONDITIONS GOVERNING PURCHASE, DISTRIBUTION, AND LIBRARY EXCHANGES.

Orders for any of the above publications should be addressed to the Division of Publications and Placement, The University of Georgia, Athens.

So long as the stock of any University publication permits, requests for single issues from interested citizens of Georgia will be honored by the Division of Publications and Placement. Libraries desiring to enter into exchange relations for The University of Georgia publications should address the University Librarian, General Library, The University of Georgia. Requests for purchase of back numbers or files of the Georgia Business Review should be addressed to the Bureau of Business Research, School of Commerce, The University of Georgia.

A COMPARISON OF THE INTELLIGENCE AND ACHIEVEMENT OF TWELVE-YEAR OLD NEGRO CHILDREN IN THE RURAL SCHOOLS OF CLARKE COUNTY, GEORGIA

By Sarah Thurmond, M. A.
Phelps-Stokes Fellow in The University of Georgia, 1931-1932

A thesis presented to the Faculty of The University of
Georgia in partial fulfillment of the requirements
for the degree of Master of Arts

CONTENTS

PAGE

PRESIDENT'S FOREWORD v

ACKNOWLEDGMENTS vi

THE PROBLEM 1

PREVIOUS STUDIES 1

NATURE OF THE STUDY 2

DESCRIPTION OF SCHOOLS SELECTED FOR STUDY 6

SUBJECTS 8

TESTS ADMINISTERED AND OTHER DATA 8

STATISTICAL TREATMENT 14

RESULTS, General Explanation of the Tables 15

AMOUNT AND RELIABILITY OF DIFFERENCES IN MEAN SCORES 19

SOME INTER-CORRELATIONS 31

SUMMARY AND CONCLUSIONS 36

BIBLIOGRAPHY 40

APPENDIX A 42

APPENDIX B 44

PRESIDENT'S FOREWORD

Upon the recommendation of the Phelps-Stokes Committee of The University of Georgia, Miss Sarah Thurmond was appointed as Phelps-Stokes Fellow for the year 1931-32. The thesis represents a study made under the direction of Dr. James Edward Greene of the College of Education.

The Phelps-Stokes Fellowship has been endowed under the following resolutions of the Trustees of the Phelps-Stokes Fund:

"Whereas, Miss Caroline Phelps-Stokes in establishing the Phelps-Stokes Fund was especially solicitous to assist in improving the condition of the Negro, and

"Whereas, It is the conviction of the Trustees that one of the best methods of forwarding this purpose is to provide means to enable southern youth of broad sympathies to make a scientific study of the Negro and his adjustment to American civilization:

"Resolved, That twelve thousand five hundred dollars ($12,500) be given to The University of Georgia for the permanent endowment of a research fellowship, on the following conditions:

"1. The University shall appoint annually* a Fellow in Sociology, for the study of the Negro. He shall pursue advanced studies under the direction of the Departments of Sociology, Economics, Education, or History, as may be determined in each case by the President. The Fellowship shall yield $500, and shall, after four years, be restricted to graduate students.

"2. Each Fellow shall prepare a paper or thesis embodying the result of his investigation, which shall be published by the University with assistance from the income of the fund, any surplus remaining being applicable to other objects incident to the main purpose of the Fellowship. A copy of these resolutions shall be incorporated in every publication issued under this foundation.

"The right to make all necessary regulations, not inconsistent with the spirit and letter of these resolutions, shall be given to the President and Faculty, but no changes in the conditions of the foundation can be made without the mutual consent of both the Regents of The University and of the Phelps-Stokes Fund."

S. V. SANFORD,
President.

---

* By the mutual consent of the Trustees of the Phelps-Stokes Fund and the Regents of The University, the Fellow in Sociology may be appointed biennially, two Fellows every three years, or annually, as the President and the Faculty may see fit. The Fellowship will then pay $750 or $500, depending upon the frequency with which it is awarded.

ACKNOWLEDGMENTS

The writer wishes to express her appreciation to the Phelps-Stokes committee of The University of Georgia, and especially to Dr. T. J. Woofter, Dean Emeritus of the Peabody School of Education. Acknowledgments are gladly given Mr. W. R. Coile, Superintendent of the Clarke County schools, and the teachers and pupils who cordially and earnestly assisted in gathering data. She is especially grateful to the director of the study, Dr. James Edward Greene, for his patient assistance and valuable suggestions and criticisms. She wishes to thank Dr. George A. Hutchinson and Dr. A. S. Edwards for suggestions and criticisms, and to express her appreciation to Mary Bryant Thurmond for encouragement.

A COMPARISON OF THE INTELLIGENCE AND ACHIEVEMENT OF TWELVE YEAR OLD NEGRO CHILDREN IN THE RURAL SCHOOLS OF OF CLARKE COUNTY, GEORGIA

I. THE PROBLEM

A comparison of the intelligence and achievement of twelve-year old Negro children in the rural schools of Clarke County, Georgia, is the problem involved in this study. In the development of this problem, objective statistical data will be offered bearing on the following topics:

(1) Raw scores made by twelve-year old rural Negro children on various standardized measures of intelligence and achievement (Appendix A).

(2) Mean scores, sigmas of distribution and sigmas of means on the several measures (Table I).

(3) The amount and reliability of differences in performance on the several measures (Table II).

(4) The inter-relationship of abilities among twelve-year old rural Negroes (Table III).

II. PREVIOUS STUDIES

A survey of investigations into state and county school systems indicates a rising demand for Negro education. The change has come gradually as the result of approximately fifty years of emancipation.

At present there are 3,379 Negro schools in the state of Georgia; thirteen of that number are rural schools in Clarke County. The Fifty-eighth and Fifty-ninth Annual School Report for the year 1929-1930 estimated the daily attendance of Negro schools to be 177,000 and the number of teachers for these pupils to be about 5,000.

School administrators, executives and educators are faced with the problem of meeting the demands for Negro education. How to deal with the Negro in his recently acquired station in society has been and still is a pertinent question put before educators and persons interested in and responsible for the welfare of the race.

To carry out programs of adjustment for the Negro, educational and psychological investigations have been made in an attempt to throw light on the characteristics of the Negro as a race.

Doctors Peterson and Lanier, (1929, p. 49) in comparing the relative abilities of sixty white and 187 Negro children in the schools of New York City on the various subtests of the Yerkes Point Scale, found "no racial difference'' on three sensory discrimination tests. Comparative scores from three memory tests on the same subjects ''favor the Negro.'' Five out of six verbal organization tests "favor the whites." In comprehension of relationships, concrete and verbal, the whites excel on three out of four tests.

Miss Ada Arlitt, (1921, pp. 179-183), tested 180 Negro children in New Orleans on the Stanford-Binet scale and, by comparing her scores with those obtained by Dr. Goddard on whites of the "same social status,'' found that the median I. Q. for Negroes decreases with increasing age from the tenth through the fifteenth year.

It is Professor Pyle's opinion that no definite statement can be made establishing superior mental abilities of whites over Negroes until standardized norms have been worked out for the Negro. The writer definitely agrees with this conclusion.

III. NATURE OF THE STUDY

The most exact method educators have for classifying an individual, group, or race is by means of an instrument of measure. This study does not attempt to compare two races and their various characteristics, but is an investigation of a group (sampling) within a race, and deals with only two traits, intelligence and achievement.

At different times the term intelligence has been interpreted and misinterpreted. Dogmatic investigators have tried to convince themselves and their colleagues that a trait such as intelligence can be accurately measured. It is well to maintain a critical attitude in the matter. "Testing persons for intellectual ability is very different from weighing cheese or measuring the length of a line with a ruler. Intelligence is not a homogeneous quantum. . . . . Individuals may be thought of as mechanisms more or less ready both by nature and by training for certain kinds of performances, varying in relative amounts from person to person. By nature some are more easily excitable along certain lines than are others; each person has special lines of aptitude and of interest. No two are quite alike.'' (Peterson, Lanier, and Walker, 1925, 271).

M. Alfred Binet, the eminent French educator, spent years in research and at various times made more or less partial statements concerning the nature of intelligence. First of all he left "faculty psychology'' and looked to the "general level'' of intelligence. (Terman, 1916, p. 43). "He (meaning M. Binet) conceives intelligence as the sum total of those thought processes which consist in mental adaptation'' (Terman, 1916, p. 43). Terman goes deeper to show that Binet's conception of intelligence broke those thought processes up into three characteristics, namely: "its tendency to take and maintain a definite direction; the capacity to make adaptations for the purpose of attaining a desired end; and the power of auto-criticism.'' (Terman, 1916, p. 45; see also Binet and Simon, 1909, pp. 1-147). In his book, Les Ideés Modernes sur les Enfants," (Binet, 1909, p. 118) Binet states: ''Comprehension, invention, direction, and criticism—intelligence is contained in these four words.''

The varying conceptions of intelligence held have not prevented attempts to design instruments of measure whereby quantitative and qualitative distinctions may be made.

Realizing the need for an impersonal and quantitative measure of school achievement educational experts have spent much time in research and experimentation in attempting to devise such measures. Objective achievement tests were the result. It cannot be maintained this early in the application of achievement tests that they are infallible; however, the coefficients of validity and reliability of such tests give the most impartial indication of their worthwhileness.

In any attempt, by research methods, to arrive at conclusions that have significance the problems must be handled in as scientific a way as possible. In making comparisons, such as are attempted in this study, if use is made of valid objective tests that have been carefully standardized on a large number of cases, one is fairly certain of accurate results. What value the tests have in respect to the function they were designed to measure depends upon the coefficients of validity that have been calculated from some criterion or criteria. For a test to be valid it must also have a high degree of statistical reliability. It is generally recognized that a given test may have a high degree of validity with respect to a given population yet not be equally valid when applied to subjects whose environment and training are radically different from that of the group on which the test was standardized.

A study of this nature requires the administration of different tests in order to obtain results that can be statistically manipulated. In some instances, a standardized test like the Binet scale may be considered an adequate measure of the intelligence of an individual or group. In another case, additional information is often needed to substantiate such ratings or to fill in gaps caused by the inadequacy of the Binet test. There are obvious reasons why the demand for various types of measures has produced the number of tests now on the market. The individual or group under investigation may not be able to understand or follow directions for a group test, requiring, therefore, an individual test score. The present writer, convinced of the inadequacy and unfairness of making comparisons based upon a single test, has based this study on material collected from two individual tests, one non-verbal, the other verbal; and five group tests, one a group intelligence test and the others, group achievement tests.

A verbal test, as the name implies, consists, for the most part, of verbal questions which require verbal answers. The

Binet scale is a widely known and used verbal test. In this test "not only are directions necessarily given in terms of language, but answers must be in the same medium. There is nothing essentially faulty in this because language is a natural overt expression of a most important type of mental activity, conveniently termed 'thinking.' " (Bigge, 1924, pp. 40-41).

Occasion arises for the use of supplementary tests that are not solely dependent upon the individual's ability to manipulate ideas. Psychologists, realizing the need, designed scales which required the subjects to "do something" rather than answer questions by word of mouth. The outgrowth of such experimentation was performance scales.

Equally different, but of another nature, from verbal and non-verbal tests are group and individual tests. Professors Levine and Marks (1928, p. 161) are of the opinion that individual tests are superior to group tests. These authors discuss the respective advantages and limitations of individual and group tests. In the first place, an individual test enables the examiner to get an insight into the mind of the subject through observing his answers, general behavior, and attitude during a testing period. Further, individual tests afford a personal, as well as a constant, stimulus that inspires a subject to put forth more thoughtful and complete responses. However efficient individual tests may be, they require a great expenditure of time. Group tests have an attractive time-saving element that is vitally important, but such tests do not bring out significant behavior trends that come from individual tests. Another advantage of group tests is the degree of objectivity they have reached, thereby, making scores obtained by different examiners highly reliable.

IV. DESCRIPTION OF SCHOOLS SELECTED FOR STUDY

Before the testing program was started Mr. W. R. Coile, Superintendent of the Clarke County Schools, was asked to rank the various rural schools for Negro children in the order of their merit. Evaluation of merit was based on building, equipment, personnel, and curriculum. Only one of the four elements remains constant, namely: curriculum.

The six best schools out of a total of thirteen were selected for intensive study. The better schools were selected partly because they had a larger number of students at the twelve year level, but mainly because it was thought that any deficiencies which might be found among the pupils of the better schools could be assumed to exist in even greater degree in the inferior schools. As accurately as Mr. Coile could discriminate among the six best schools (out of a total of thirteen), they come in order of merit as follows:

(1) Teacher Training and Industrial Institute
(2) Midway
(3) St. James
(4) St. Luke
(5) Morton's Chapel
(6) Chestnut Grove.

(1) Teacher Training and Industrial Institute, six miles from Athens on the Danielsville highway, is a brick structure. The main building includes three teaching rooms, an auditorium, office, and library. A building on the grounds affords vocational training laboratories. This is the only school in the list selected that had any adequate equipment. Here the desks, of the patented variety, are partially suited to the size of the pupil. There are maps and fairly good blackboard facilities. The rooms are better lighted and ventilated than heated. However, in spite of the limitations of the building itself, the teachers have tried to make the rooms as attractive as possible.

(2) Midway, located on the Athens-Winterville road, is a two-room, frame building. Both rooms are fairly large, but heating apparatus is in poor condition. Lighting and ventilation are fair. Here there has been no small amount of time

and ingenuity spent to make the building and grounds more attractive. The equipment is scanty, and of poor quality. Desks are made of rough timber, not designed to fit the size of the prospective occupant, thereby making seating a tiresome and ofttimes painful process. The blackboard space is slight, and not in good repair. Teachers furnish the material with which they have to work, such as chalk and erasers.

(3) St. James, approximately four miles from Athens on the Gainesville highway, is a two-story lodge-hall school combination. A dark, dangerous stairway connects the two school rooms. Both rooms are in need of repair, especially the ground floor room which has an unsanitary and unhealthy drainage system. Neither room is large enough to accommodate the number of children in regular attendance. The seating equipment is the same as that described in Midway school. Blackboard space here is limited to rough planks smoothed enough to take a coat of black paint. The paint is wearing and the planks no longer serve as an adequate excuse for a blackboard.

(4) St. Luke, located on the Athens-Augusta road about five miles from Athens, is a two-story, two-room, frame structure. The desks are the rough type described in an earlier statement. The teachers have to furnish what equipment they need, such as crayon and erasers. In this school as well as the others visited the teachers have tried to make the building and grounds attractive, however dilapidated the conditions of the buildings are.

(5) Morton's Chapel is located on the Barnett Shoals road about eight miles from Athens. This is a one-teacher, one-room school. The room is large enough for one teacher but is not sufficient to accommodate all the grades. The building is in need of repair. There is very little variation in building and equipment among any of the schools with the exception of the Training school. However, the teacher has tried, under the prevailing conditions, to make the room livable.

(6) Chestnut Grove, situated on the Epps' Bridge road about seven miles from Athens, is the smallest school in size, enrollment, and building of all the schools visited. It is

draughty, poorly lighted, and poorly heated. One teacher is in charge of all the grades from the first through the seventh. The desks, blackboard, and miscellaneous equipment are in poor repair, and fail to give even average service.

V. SUBJECTS

For this study a list was made of every pupil in the six rural schools selected who had reached the age of twelve, but had not passed the age of thirteen.

Children of this age were selected because it is thought to be the most representative of the degrees of mental ability found in the schools. All the dull children who are not wholly deficient will be enrolled and the bright ones will not have left the grades (Peterson and Lanier, 1929, p. 6). Rural subjects were selected partly because the writer is more familiar with the rural Negro than with the urban Negro, and partly because of the greater need for research among rural Negroes.

Data were secured on the various measures from fifty-two twelve-year old Negro children. On account of absences from school, illnesses, and other ''providential'' reasons existing at the time of testing, only forty of this number furnished complete data. All results are therefore based on forty subjects.

VI. TESTS ADMINISTERED AND OTHER DATA

PROCEDURE. Mental testing depends upon the examiner's familiarity with the testing materials and ease in handling the subject and the situation. To insure accuracy in test results the director of this study, Dr. James Edward Greene, instructed the writer in the administration and scoring of each test chosen for the study. As a precautionary measure he, at various times, observed the testing during the collection of data from the field.

When sufficient knowledge of how to administer and score each measure was acquired, the investigator began the collec-lection of data. The individual tests (Arthur and Stanford) were given in the order mentioned. In administering the Arthur test, directions set down by Dr. Grace Arthur in her manual (1930) were followed verbatim. The problem of securing adequate testing space was one that had to be met with the aid of the teachers of the various schools. Due to the lack of available space in the schools for the administration of the individual tests, the teachers in some of the schools arranged for the writer to use the church (when the two were on the same lot). Far from an ideal place, but one that was familiar to the child, and one that afforded complete quiet, the church served exceptionally well. The church was used at St. James and Chestnut Grove. At the Teacher Training and Industrial Institute, the library was turned over to the investigator.

When the weather was sufficiently warm to enable testing out of doors such was the program, but in case the cold was severe the teacher arranged for the investigator to use one of the two class rooms. At Morton's Chapel the weather permitted the writer to administer all of the individual tests out of doors and the same procedure was followed at St. Luke's and most of the time at Midway.

The testing began in October and was completed in May. Caution was taken to finish a child's entire testing program within the age first recorded, so that an irregularity of months would not complicate final scoring. In the case of the performance test, time is an important element in scoring. In keeping with customary procedure, a slide stop watch was used in recording all timed tests.

The Stanford test was administered as the second test in the series. Suggestions for giving and scoring the measures given in the manual (Terman, 1916) were followed to the letter. Accurate time was kept on this test when required. Verbatim responses to the various test situations were recorded and are on file for inspection. The same facilities as to space, etc., were present for this test as for the performance test.

The administration of the group tests (Illinois, reading, arithmetic, spelling and handwriting) were given in the order mentioned. Directions for administration and scoring suggested by the respective authors were followed implicitly. No variation occurred from one day to the other, or from one group to the other. On all tests requiring accurate notation of time the slide stop watch was used. Arrangements were made beforehand with the individual teachers to provide space to give the group tests. A separate room, not entirely free from distractions but as nearly free as the usual school room, was used in each school for the administration of each group test. Regulation lined copy paper, and pen and ink were furnished by the writer for the handwriting test. Paper was also furnished for the spelling test.

As a precautionary measure each test score was double checked for errors in scoring, and the results are on file for inspection.

DESCRIPTION OF THE TESTS USED. Following the general purpose of this study, intelligence and achievement tests were given in order that group comparisons on the two functions might be made. For this investigation the following measures were used: two individual tests of intelligence (Arthur and Stanford); one group test of intelligence (Illinois); four group achievement tests (reading, arithmetic, spelling and handwriting). The following is a description and the order of administration of each measure in the series:

(1) Arthur Point Performance Scale (which we shall call Arthur) is an individual intelligence test of the non-verbal type. According to the author (1930, p. 1, ff.) this test is designed to ''support the Binet rating;'' ''to serve as a shock absorber;'' "to provide an opportunity to make a satisfactory contact with the patient,'' and "to observe the behavior of patients in a variety of situations;'' as a measure of the ability to "compete successfully'' with persons of one's own age in a non-school situation; "to give an adequate report of the intelligence of the individual in whom verbal and non-verbal abilities are markedly unequal in their development;" as a substitute scale "where a valid Binet rating cannot be obtained because of language difficulty, speech defect or emotional repression;" and, as "a valuable corrective to the Binet rating.'' In the latter case, Arthur states, (1930, p. 7), "the Binet rating tends to reflect the environment, while the performance rating gives a safer index of the ability of the child to meet new situations.'' The test is described, evaluated, and interpreted

by the author in her manual *A Point Scale of Performance Tests* (Arthur, 1930). The time required for giving this test to twelve-year old rural Negro children varies from forty minutes to two hours.

Using conformity with Binet I. Q.'s as a criterion of validity the writer computed (from data furnished by Arthur*) the average probable error of measurement to be 5.82 I. Q. points in subjects ranging in mental age from five through fifteen years. The same author (1930, p. 5) reports a P. E. of measurement of 6.19 I. Q. points for fifty-two twelve-year old subjects between Form I and Form II when retested approximately one year later.

(2) Stanford Revision of the Binet-Simon test (which we shall hereafter designate as Stanford) is one of the numerous revisions of the Binet-Simon general intelligence scale, and perhaps the most widely known and used test of this nature.** This scale "does not pretend to measure the entire mentality of the subject, but only *general intelligence*'' (Terman, 1916), p. 48). By implication at least Terman (1916, pp. 230, 344, etc.) gives the impression that the Stanford is a test of verbal intelligence. He states, "it is the very essence of the higher thought processses to be conceptual and abstract.'' (1916, p. 344).

Directions, as given by Terman in *The Measurement of Intelligence,* for administration and scoring were followed. The time for each individual test was on the average somewhat greater than that required for the performance scale. For an evaluation and interpretation of each test in the scale consult the manual.

Dr. Terman found a reliability coefficient (calculated from scores of 428 subjects from the third to fifteenth age level) to be .93. Professor W. M. Proctor found a validity coefficient of .74 on scores of 116 adults that had Army Alpha scores.

(3) Illinois Examination, General Intelligence (called merely Illinois for brevity) is a group intelligence test. For the most part this is the first test administered to our sub-

---

* Data given in a personal letter to the writer.

** Stanford is often taken as a reliable criterion of general intelligence. (Garrett, 1931, p. 266).

jects which required any special ability to read, and to follow written directions. Form 2, designed for grades three, four, and five, was used. This measure of intelligence was made up of the following subtests: analogies, arithmetic problems, sentence vocabulary, substitutions, verbal ingenuity, and synonym-antonym. The authors claim that the test is a measure of "general intelligence or general mental ability" and state that " in a school situation we may think of the intelligence of a pupil as his inherent capacity to learn or do the usual work of the school; it might be termed his 'schoolability'." (Handbook, 1920, p. 11).

Administration and scoring routine as set out in the Handbook (1920) was followed by the investigator. Results on the two forms, 1 and 2, were correlated by the authors who found a reliability coefficient of .92 and a P. E. of measurement of 5.3 I. Q. points. The Illinois scale was given to 203 pupils whose mental ages had been determined by Stanford. The correlation between the mental ages determined by these two scales is .74 ± .02 (Handbook, 1920, p. 31).

(4) Monroe's Standardized Silent Reading Test (Revised) of the Illinois Examination is a group test designed to measure the rate and comprehension of silent reading. The time limit on this test is exactly four minutes, and no subject finished before time was called. Consult the Handbook for a full description and evaluation of the measure. The authors report reliability coefficients as follows: Rate, .78; Comprehensions, .80 (Handbook, 1920, p. 31).

(5) Monroe's General Survey Scale in Arithmetic of the Illinois Examination, designed to measure arithmetic achievement for grades three, four, and five, was used. Arithmetic problems ranging from simple to difficult, and embracing the fundamental processes, comprise the test. Each subject has a short time limit, ranging from thirty seconds to three minutes. Directions are set down in the Handbook explaining the method of administration and scoring. A reliability coefficient of .95 was calculated by the author of the test for grades three, four, and five. Using cases from grades three

through eight, Ruch and Stoddard found a correlation of .76 between silent reading and arithmetic.

(6) Ayer's Spelling Scale, a group-test of spelling achievement, was given to every subject. The purpose of this scale is to measure an individual's spelling ability. In an adaptation of this test by Dr. Marion Monroe (1930, pp. 3-4) the words are arranged in groups of ten words to a grade level. In this study, words for grades one through eight were used. The test provides norms for both written and oral spelling. The raw scores were converted into B-Score (i. e., grade-placement norms), according to directions furnished by Dr. Monroe (Manual, 1930, p. 4), and then translated into age norms in order to secure uniformity of data (scores on all the measures are stated in terms of months). In converting B-Scores into age scores, use was made of the age-grade distribution (based on 501,521 school children) worked out by the Research Division of the University of California. (Guide for Interpreting New Stanford Achievement Test, 1929, p. 10).

(7) Ayer's Handwriting Scale is a test designed to measure the rate and quality of an individual's handwriting. The "Gettysburg Edition'' was used in this study. Furnished instructions as to administration and scoring were rigidly adhered to. The time required for the test proper was exactly two minutes. It took about ten minutes to prepare and instruct the subjects, and permit them to become accustomed to the pen staff and ink. Arthur S. Otis (1925, p. 755) correlated results from this scale with those on Starch's Handwriting Scale and found the coefficient to be .876. Only the scores on rate of writing lent themselves to conversion into age norms from the B-Score norms. The scores on quality of writing are, therefore, excluded from this study. In converting the B-Scores on speed of writing into age scores use was again made of the age-grade distribution furnished by the University of California.

VII. STATISTICAL TREATMENT

All the original data are on file and are open to inspection. Limitation of space makes it inadvisable to present here the original data in their entirety. For the benefit of readers who may desire it, a table giving the most essential data (mental ages, educational ages, intelligence quotients, educational quotients, chronological age, sex, grade placement) is furnished. All this information will be found in Appendix A.

The following additional statement concerning statistical treatment may be in order here.

Measure of Central Tendency. The mean was used as the measure of central tendency. The mean was found by the "Short method" (Garrett, 1931, p. 53). Table I gives the mean scores on the several measures.

Measure of Variability. The standard deviation is used and is derived from the formula for grouped data (Garrett, 1931, p. 52). Table I gives this information.

Measures of Reliability. The reliability of the mean in terms of its standard deviation was calculated from the usual formula (Garrett, 1931, p. 121). The reliability of differences between two measures was calculated from a formula given by Garrett (1931, p. 129). Table I furnishes data on the reliability of means while Table II gives the measures of reliability of differences in means.

Measure of Relationship. The measure of relationship between variables was calculated from Professor Karl Pearson's formula for finding the coefficient of correlation **r** by the product-moment formula: .
Table III gives the correlation between each measure and each other measure with their respective probable errors. The probable error of **r** was obtained from a statistical table prepared by Karl J. Holzinger. (1925, pp. 60-69).

Scoring. All mental and educational ages were worked out according to directions given in various manuals and handbooks. I. Q.'s were calculated in the conventional way following the formula: I.Q.=M.A./C.A. The educational quotients in arithmetic and reading were calculated from the

TABLE I.

MEAN SCORES, SIGMAS OF DISTRIBUTION, AND SIGMAS OF MEANS OF EACH OF THE SEVERAL MEASURES

Measures	Mean Scores	Sigma of Distribution	Sigma of Means
Chronological Age	149.95	4.23	.67
Arthur M. A.	125.50	19.65	3.10
Stanford-Binet M.A.	114.50	14.27	2.25
Illinois M. A.	100.13	18.97	2.99
Spelling E. A.	107.33	17.58	2.77
Arithmetic E. A.	99.50	15.16	2.39
Reading Comprehension E. A.	92.38	12.09	1.90
Reading Rate E. A.	89.25	11.32	1.78
Handwriting E. A.	87.38	15.10	2.27
Arthur I. Q.	.84	12.28	1.92
Stanford-Binet I. Q.	.77	8.37	1.32
Illinois I. Q.	.69	12.40	1.97
Spelling E. Q.	.71	11.10	1.75
Arithmetic E. Q.	.66	10.50	1.50
Reading Comprehension E. Q.	.62	7.80	1.23
Reading Rate E. Q.	.61	6.80	1.07
Handwriting E. Q.	.59	9.54	1.50
Spelling A. Q.	.94	11.48	1.81
Arithmetic A. Q.	.87	17.45	2.75
Reading Comprehension A. Q.	.82	9.51	1.50
Reading Rate A. Q.	.80	9.13	1.44
Handwriting A. Q.	.77	6.42	1.01

Handbook of the Illinois Examination. The A. Q.'s for educational achievement tests were calculated by using the Stanford I. Q. and the various E. Q.'s previously obtained. From the formula A.Q.=E.Q/I.Q. The Stanford I. Q. was used as the constant in these calculations for the reason that given Stanford mental ages are generally accepted as being reasonably valid measures of ability to do formal school work of a given difficulty.

VIII. RESULTS

GENERAL EXPLANATION OF THE TABLES

In a consideration of the results of this study, effort is made to present the data in such a form as to make easy the direct comparison of scores on different tests. Since all scores on the various measures are stated in terms of months such a direct comparison may be easily made. For convenience and clearness the general plan will be followed in the inter-test

TABLE II.

Reliability of Differences in Mean Scores of the Several Measures N 40.

Measures Compared	Difference in Means	Sigma Difference in Means	** "Critical Ratio"	*** Chances of Reliability
C. A.—Arthur M.A.	24.45	3.17	7.71	Comp. Rel.
C. A.—Stanford M. A.	35.45	2.35	15.08	Comp. Rel.
C. A.—Illinois M. A.	49.82	3.06	16.28	Comp. Rel.
C. A.—Spelling E. A.	42.62	2.85	14.95	Comp. Rel.
C. A.—Arithmetic E. A.	50.45	2.48	24.41	Comp. Rel.
C. A.—Reading Comprehension E. A.	57.57	2.02	28.50	Comp. Rel.
C. A.—Reading Rate E. A.	60.70	1.90	31.90	Comp. Rel.
C. A.—Handwriting E. A.	62.57	2.36	26.51	Comp. Rel.
Arthur M. A.—Stanford M.A.	11.00	3.83	2.87	99.74
Arthur M. A.—Illinois M. A.	25.37	4.31	5.88	Comp. Rel.
Arthur M. A.—Spelling E. A.	18.17	4.16	4.37	Comp. Rel.
Arthur M. A.—Arithmetic E. A.	26.00	3.91	6.65	Comp. Rel.
Arthur M. A.—Reading Comprehension E. A.	33.12	3.64	9.09	Comp. Rel.
Arthur M. A.—Reading Rate E. A.	36.25	3.57	10.15	Comp. Rel.
Arthur M. A.—Handwriting E. A.	38.12	3.84	9.92	Comp. Rel.
Stanford M. A.—Illinois M. A.	14.37	3.74	3.84	Comp. Rel.
Stanford M. A.—Spelling E. A.	7.17	3.57	2.00	98.
Stanford M. A.—Arithmetic E. A.	15.00	3.27	4.59	Comp. Rel.

TABLE II.—(Continued)

RELIABILIRY OF DIFFERENCES IN MEAN SCORES ON THE SEVERAL MEASURES N 40.

Measures Compared	Difference in Means	Sigma Difference in Means	** "Critical Ratio"	*** Chances of Reliability
Stanford M. A.—Reading Comprehension E. A.	22.12	2.94	7.52	Comp. Rel.
Stanford M. A.—Reading Rate E. A.	25.25	2.87	8.76	Comp. Rel.
Stanford M. A.—Handwriting E. A.	27.12	3.19	8.50	Comp. Rel.
Illinois M. A.—Spelling E. A.	— 7.20*	4.07	—1.77*	96.
Illinois M. A.—Arithmetic E. A.	.63	3.83	.16	56.
Illinois M. A.—Reading Comprehension E. A.	7.75	3.54	2.19	99.
Illinois M. A.—Reading Rate E. A.	10.88	3.46	3.14	Comp. Rel.
Illinois M. A.—Handwriting E. A.	12.75	3.75	3.40	Comp. Rel.
Reading Rate E. A.—Handwriting E. A.	1.87	2.88	.65	74.
Reading Comprehension E. A.—Reading Rate E. A.	3.13	2.60	1.20	88.
Reading Comprehension E. A.—Handwriting E. A.	5.00	2.96	1.68	94.
Arithmetic E. A.—Reading Comprehension E. A.	7.12	3.05	2.33	98.90
Spelling E. A.—Arithmetic E. A.	7.83	3.65	2.14	98.
Arithmetic E. A.—Reading Rate E. A.	10.25	2.96	3.46	Comp. Rel.
Arithmetic E. A.—Handwriting E. A.	12.12	3.29	3.68	Comp. Rel.
Spelling E. A.—Reading Comprehension E. A.	14.95	3.36	4.45	Comp. Rel.
Spelling E. A.—Reading Rate E. A.	18.08	3.29	5.49	Comp. Rel.
Spelling E. A.—Handwriting E. A.	19.95	3.58	5.57	Comp. Rel.

* Minus difference indicates that the difference favors the second named variable.

** "Critical Ratio" is derived from: Difference in Means/Sigma Difference in Means.

*** "Chances of Reliability" means that there are so many chances in 100 that a true difference greater than zero exists. "Comp. Rel." indicates that the difference found is a completely reliable difference.

comparisons, of naming first the test on which the higher score is made.

The *size* of the difference in means and its reliability ("critical ratio") for each of the inter-test comparisons is given in Table II. Table I furnishes the mean score, sigma of distribution, and sigma of means for each of the measures. Although our discussion of the findings shall mainly be confined to *ages* rather than *quotients*, we have included data on quotients in Table I for the benefit of the reader who may prefer to think in terms of quotients rather than of ages.

Table II gives the difference in means, the sigma of difference in means, and the "critical ratio." The last named measure indicates the degree of probability that the difference found between any two tests is not due to chance factors. The reader should bear in mind that differences are not considered reliable if the "critical ratio" given in Table II (difference in means / sigma of difference in means) is less than three. Since a "critical ratio" of as much as three is to be taken as an indication that differences greater than zero exist, it should be noted that a "critical ratio" beyond three may be taken as indicating just so much added reliability. (Garrett, 1931, p. 133). That unreliable differences are found is not to be taken as evidence that such differences do not actually exist or that they are certainly due to chance factors. Data based on too few cases may fail to show statistically reliable differences despite the presence of a general tendency for such differences to exist.

It should also be pointed out that even when differences are shown to be statistically reliable, caution should be used in the interpretation of these differences. Statistically reliable differences in means of two different racial groups, or on two measures of the same group, may be due to inadequacies of measurement, to improper motivation, or to differences in educational and cultural status rather than to differences in innate ability. It is necessary to proceed slowly and cautiously in the making of interpretations derived from statistical comparisons, for hasty conclusions tend to obstruct scientific progress by bringing about suspicions, bias, failure to cooperate, etc. With imperfect data conclusions must be regarded as tentative.

AMOUNT AND RELIABILITY OF DIFFERENCES IN MEAN SCORES

Since it is the primary purpose of this study to make detailed comparisons of the intelligence and achievement (using standardized tests of intelligence and achievement as criteria) of twelve-year old rural Negro children the following order of presentation of data will be adhered to:

1. A comparison of chronological age with the various mental ages and educational ages. This data will indicate the amount and reliability of mental and educational retardation or acceleration.
2. A comparison of each mental age with each educational age. This data will indicate the amount and reliability of retardation or acceleration in school attainment as judged by each of the intelligence tests.
3. A comparison of each mental age with each other mental age. Here we shall have a comparison of Negro mentality on the basis of the three tests of intelligence.
4. Comparison of each educational age with each other educational age. This data will indicate the relative efficiency of twelve-year old Negroes in the various school subjects under the conditions of this study in so far as educational tests measure educational efficiency.

In making comparisons of the sort indicated above it is necessary to show the *amount* as well as the *reliability* of any difference found. It is, of course, obvious that a considerable difference may exist without that difference being reliable It is also true that a difference may be reliable and yet be small. A true picture of the situation, therefore can be obtained only if both of these factors are taken into account. In our discussion we shall not attempt to explain or interpret any differences found until we have shown the degree to which those differences are statistically reliable. We shall, therefore, first of all give a general discussion of the reliability of our differences.

General Statement of Reliability of Differences

Of the various comparisons made in Table II, twenty-six out of thirty-six differences in means are completely reliable. It should be noted that of these ten incompletely reliable differences, all except three (Illinois—arithmetic, reading-rate-reading comprehension, and reading rate-handwriting) have critical ratios sufficiently high as to indicate virtually complete reliability. The chances in 100 that a true difference greater than zero exists are indicated in the last column of Table II. The differences between chronological age and all three of the measures of intelligence (Arthur, Stanford, and Illinois) are more than completely reliable. A comparison of chronological age with achievement age shows that all the differences are completely and highly reliable.

A comparison of Arthur mental age with the various educational ages shows completely reliable differences in all comparisons. However, in this case the lowest critical ratio (4.37) is found between Arthur and spelling. The highest critical ratio (10.15) is between Arthur and reading rate. Stanford scores, when compared with the various achievement scores, show completely reliable differences favoring Stanford in all cases except spelling. In the case of Stanford and spelling the chances are ninety-eight out of 100 that the difference is reliable. The most completely reliable differences between Stanford and any of the achievement measures is found in reading rate (critical ratio, 8.76). Differences between the Illinois test and the various achievement measures are completely reliable and favor Illinois in the case of reading rate (critical ratio, 3.14) and handwriting (critical ratio, 3.40). Note should also be taken of the following comparisons: Illinois with reading comprehension the chances are ninety-nine out of 100 that the difference (favoring Illinois) is completely reliable; Illinois with arithmetic the chances are fifty-six out of 100 that the difference (favoring Illinois) is completely reliable; Illinois with spelling the chances are ninety-six out of 100 that the difference is reliable (difference favors spelling).

A comparison of the differences between each intelligence measure and each other such measure shows that the differences are completely reliable between Arthur and Illinois and between Stanford and Illinois; Arthur and Stanford give a high but not completely reliable difference (99.74 chances out of 100) in favor of Arthur.

Inter-test comparisons on the various achievement measures show that completely reliable differences exist in five out of ten comparisons (completely reliable superiority in spelling over handwriting, reading rate, and reading comprehension; completely reliable superiority in arithmetic over reading rate and handwriting). Incompletely reliable differences exist on the following comparisons: reading comprehension over reading rate (critical ratio, 1.20); reading comprehension over handwriting (critical ratio, 1.68); reading rate over handwriting (critical ratio, .65); spelling over arithmetic (critical ratio, 2.14); arithmetic over reading comprehension (critical ratio, 2.33).

It is thus seen that our subjects (as measured by these tests) are reliably most proficient in spelling and arithmetic and least proficient in handwriting.

Now that we have given a general statement concerning the extent to which the differences here found are reliable we shall proceed to discuss in detail the amounts of difference found in the various comparisons and to attempt an explanation of our findings.

Mental and Educational Retardation Estimated From Chronological Age

A comparison of the differences in mean scores between chronological age and both the intelligence and achievement scores shows that the subjects under study do markedly, and in the main reliably, less well on all measures of intelligence and achievement than do other non-selected twelve-year old children on whom the norms for the several tests are based.

The differences found from comparing the mean scores on chronological age and mental age follow: Between chronological age and Arthur mental age, 24.45 months (approximately two years mental retardation); between chronological age and

Stanford mental age, 35.45 months (almost three years mental retardation); between chronological age and Illinois mental age, 49.82 months (approximately four years mental retardation).

Accepting a difference between chronological age and the mental ages here measured as an index of retardation the differences indicate that our subjects are markedly (and virtually reliably) less retarded mentally on Arthur than on Stanford. They are greatly and reliably more retarded on Illinois than on Arthur. The samplings are markedly and reliably more retarded on the Illinois scale than on either Arthur or Stanford.

In accounting for the retardation (two to four years) on the measures of intelligence it must be remembered that, as yet, no separate norms for rural Negro children have been prepared on these tests. The scores were based on norms prepared from scores made by children of different race, social level, home environment, and school training. Consequently, the scores made by our subjects serve merely as a basis for comparing their performances with the performances of other non-selected twelve-year old children rather than with the average performance of twelve-year old Negroes. Another factor of possible significance is the conditions under which the testing was made. The accommodations (over which there was no control) for administering the tests were of such a nature that they may have affected the subject's responses to some degree. The nature and extent of such influence, if present, cannot be determined. Although our data indicate marked mental retardation, it could reasonably be assumed that, if separate racial norms on each of the tests were available and adequate provisions were made for administration of all tests, the subjects would be less retarded.

Factors of inadequate norms and administration do not, however, entirely explain the enormous retardation here found. Treated in a non-objective manner, the extent to which home environment, social status, irregular school attendance and low cultural background influence the child's individual mental ability is not known. However, it is apparent from daily observation as well as from scientific studies (see Freeman, et. al., 1928) that a child from a home of average or high cultural environment is decidedly more mentally alert (if physically normal) than a child from a low cultural environment. The school situation is generally conceded to have a direct effect upon the subject's responses on both the Stanford and Illinois tests. It is noted that for our subjects the amount of mental retardation is greater on these two measures than on Arthur. To what extent the factors of irregular attendance, poorly equipped schools, lack of drill, ineffective teaching or grade placement, influence the scores on these tests cannot be said unless all other contributing factors were controlled. Nevertheless it is perhaps safe to assume that the school situation is in part responsible for the retardation on Stanford and on Illinois. Whatever the factors may be which account for the amount of retardation as well as the varying degrees of retardation here found as estimated from the three measures of intelligence, we are perhaps safe in assuming that under existing conditions twelve-year old rural Negroes show decidedly more aptitude for non-verbal performances than they do for verbal activities.

When the various measures of educational achievement are compared with chronological age the following amounts of educational retardation are noted: 60.70 months in reading rate; 57.57 months in reading comprehension; 50.45 months in arithmetic; 42.62 months in spelling and 62.57 months in handwriting. The amount of educational retardation as estimated from chronological age ranges from 42.62 months (three and one-half years) on spelling to 62.57 months (more than five years) on handwriting. All of the above differences are enormous and very highly reliable (critical ratios range from 14.95 to 26.51).

In accounting for the failure of our subjects to succeed on school material that is especially designed to measure achievement one must look to the home environment and the school situation as contributing factors. How the home environment influences the child's ability on various school subjects must of necessity be handled in a subjective manner. It is the opinion of the writer that conditions found in the average home of rural Negroes cause irregular attendance. An inspection of numerous homes from which the subjects came shows in some instances that the houses are small and inadequate for the housing of large families. There is little food, especially in the homes of croppers and sharers. The food is also inadequate from the standpoint of balanced diet, etc. There is little evidence of "cultural" tastes, according to white standards. From such homes the majority of rural Negro children come. There are exceptions, however, where the Negro owns the farm, and has some adequate means of housing and supporting a family.

In consort with the home situation is the school. The amount of influence the school has on educational retardation cannot be estimated in an objective way. However, it is the purpose of the investigator to bring out some significant characteristics of the school room and the method of teaching in an effort to arrive at adequate conclusions. The dilapidated condition of the average rural Negro school, the crowded conditions, the lack of books and regular attendance, the inadequate methods of teaching and the poorly trained teachers are all significant contributing factors. It is the opinion of the writer that the most influential factor involved is that of school attendance. It should be emphasized that the attendance law is not rigidly enforced upon the Negro. Regular attendance will undoubtedly expose the pupil to the subjects taught in the school curriculum. In spite of the fact that there is very little individual attention given the pupils it stands to reason that a slightly higher degree of mastery of school subjects would follow regular attendance.* The reason for irregular attendance is evident from observing the social and economic standing of the parents. A goodly number of children do not enter school until the harvest season is over and they are taken out of school for spring planting. Another aspect of the home environment in a good many cases also influences attendance. An example of this type of home factor is exemplified in the case of a parent disliking the teacher.

---

* It has been utterly impossible to secure records of attendance except for sixteen of our subjects.

Another instance is the child's dislike for the teacher. Animosity among the school children causes some of the younger children to stay at home unless they have an older brother or sister to take care of them on the school grounds.

Grade placement unquestionably influences the pupil's inability on the various school subjects. Over-placement, as measured by standardized achievement tests, is particularly noticeable in the rural Negro schools, and is conducive to discouragement, lack of application, low morale, etc. The extent to which our subjects are under-placed, correctly placed or over-placed (as estimated by the various measures of intelligence and achievement) is indicated in Appendix B.

EDUCATIONAL RETARDATION AS ESTIMATED FROM THE VARIOUS MEASURES OF INTELLIGENCE

The differences in mean scores between Arthur mental ages and the various achievement ages indicate the following amounts of educational retardation: spelling, 18.17 months; arithmetic, 26 months; reading comprehension, 33.12 months; reading rate, 36.25 months, and handwriting, 38.12 months. The amount of retardation ranges from 18.17 months (more than one and one-half years) on spelling to 38.12 months (more than three years) on handwriting. The differences are not only large but completely reliable.

Using Stanford scores as a criterion of mental ability it is found that the differences (favoring Stanford) are as follows: spelling, 7.17 months; arithmetic, 15 months; reading comprehension, 22.12 months; reading rate, 25.25 months; handwriting, 27.12 months. The amount of retardation ranges from 7.17 months to 27.12 months (more than two years). All the differences except spelling are significantly large as well as reliable. While all of the above differences favor Stanford, they are, in no instance, either as large or as reliable as the differences between Arthur and the same educational scores.

If Illinois scores are used as a basis for comparing intelligence with the various achievement measures we find the following amounts of educational retardation: Arithmetic, .63 months; reading comprehension, 7.75 months; reading rate, 10.88 months, and handwriting, 12.75 months. An incompletely reliable superiority in spelling over Illinois of 7.20 months is to be noted. Reading rate and handwriting give the highest (completely reliable) positive differences. The amount of educational retardation, estimated from this criterion, ranges from .63 months (arithmetic) to 12.75 months (handwriting). All of the above differences are relatively small and only two of them (reading rate and handwriting) are completely reliable. While all of the above differences except spelling favor Illinois, they are, in no instance, either as large or as reliable as the differences between Arthur and the same educational measures. They are also consistently smaller and less reliable than the differences found between Stanford and the various educational measures.

To summarize the data on educational retardation as estimated by the three measures of intelligence: Our subjects show a very much greater amount and reliability of educational retardation when Arthur (all five comparisons show enormous and very highly reliable differences) is used as a criterion of ability to master school subjects, than when either Stanford or Illinois is used as a criterion. The amount of retardation as estimated from Stanford is both greater and more reliable than when estimated from Illinois. However, when Illinois is used as a criterion the amount of retardation is significantly great, on all subjects except spelling, and, in the main, reliable. We are, therefore, forced to conclude that twelve-year old rural Negro children, under the conditions of this study, have the ability to succeed in school work to a much greater extent than has been done. The wide discrepencies in the amount of educational retardation as estimated by the three measures of intelligence may fairly be taken as evidence that the intelligence tests in questions are not equally valid measures of the ability of our subjects to do formal school work. A tentative evaluation of the validity, for our subjects, of these three tests of intelligence follows.

TENTATIVE EVALUATION OF THE VALIDITY, FOR OUR SUBJECTS, OF THE THREE MEASURES OF INTELLIGENCE

If the differences in means of the three measures of general intelligence are compared it is found that the amounts of mental age difference are as follows: between Arthur and Stanford, 11 months in favor of Arthur; between Arthur and Illinois, 25.37 months in favor of Arthur; between Stanford and Illinois, 14.37 months in favor of Stanford. These differences may be regarded as significantly large since they range from 11 months to 25.37 months. The differences between Arthur and Illinois and between Stanford and Illinois are completely reliable, while the differences between Stanford and Arthur are virtually reliable (critical ratio, 2.87).

Therefore, it is apparent from the data presented that our subjects show more marked proficiency in manipulating concrete objects than in dealing with abstract ideas or verbalizations.

Other investigators, Peterson and Lanier, (1929), Arlitt, (1921), and Pyle, (1915), mention the low cultural background of the Negro as a possible explanation of his varying attainments on mental tests. That social status is a significant factor in determining one's rating on mental tests is generally recognized. Separate norms for rural Negro children on the tests used are not available. It can be safely admitted, however, that rural Negro children are placed at a disadvantage because of cultural differences if they are compared with a non-selected group representative of the total population. Whether the factor of norms entirely accounts for the low mean scores is a question yet to be answered by further research in the field.

There is evidence at hand to show that instead of merely diagnosing differences in innate ability many mental tests tend to diagnose, or at least to reflect strongly in their diagnoses, differences in educational and cultural status. "The relative superiority of school children over adults could scarcely be accounted for on the basis of differences in their innate abilities, but must probably be looked for in differences in the immediate environment.'' (Thomas and Thomas, 1928, p. 338).

Isserlis, working with Burt's tests on English school children, found "a distinct correlation between intelligence of school children and their environment, whether measured by the economic position of the parents, by the care taken of the home, or by the clothing of the children.'' (Isserlis, 1923, pp. 17-10). Freeman, Holzinger, and Mitchell, (1928) found an appreciable and statistically reliable gain on Stanford I. Q.'s in seventy-four cases of children who had spent considerable time in a foster home of higher cultural status than their own homes. Data of this sort seem definitely to indicate that scores on such scales as the Stanford and Illinois are influenced by cultural and academic status.

It is not difficult to account for the differences which favor Arthur. The test apparently does not depend upon a high cultural or school training background Considering the importance of cultural background and school training in their relation to Stanford and Illinois it is easily understood that that measure which is least dependent upon these factors would yield a higher score from twelve-year old negro subjects whose school training and cultural status are known to be inferior.

The wide variation in the amount of mental retardation here found, as estimated from the three measures of mental ability, suggests either (a) that the three tests should not be regarded as measures of the same thing—presumably "general intelligence''—but that each measures a specialized "type'' or "aspect'' of intelligence; or (b) that, if they measure the same thing, they do so unequally well, due to the influence of cultural factors. It is the writer's opinion, admittedly incapable of adequate proof from the limited data at hand,* that both of the above factors are operative. She regards the Arthur test as being mainly a measure of general adaptability or of "non-verbal'' intelligence, and as such, relatively little influenced by cultural factors. The Stanford is regarded as being more largely a measure of a somewhat specialized (verbal) ability, i. e., the ability to carry on abstract and conceptual thinking of a sort usually required in a school situation. Since abstract thinking is a verbal phenomenon, and since language development is greatly dependent upon cultural status, and since the cultural status of our subjects is known to be inferior, as judged by "white" standards, it is felt that the Stanford test is a less valid measure of the innate capacity of our subjects than is the Arthur test. The Illinois test is regarded as measuring much the same ability as does Stanford only it does so much less adequately for our subjects for the reason that it is much more greatly dependent upon formal school training, and to some extent more dependent upon cultural factors of a non-school sort. The tentative hypothesis is advanced that the Arthur test is the most valid test of the innate capacities of our subjects, whereas the Illinois is the least valid such measure.

---

* This opinion is based on a detailed analysis of supplementary data which, because of its voluminousness and statistical unreliability is unsuited for presentation here. It is hoped, however, that this data may be supplemented and written up as an article for one of the technical journals at a later date .

Relative Degrees of Educational Proficiency

If we compare the mean scores on the various achievement tests the following differences (differences favor the first named variable) are noted: reading rate—handwriting, 1.87 months; reading comprehension—reading rate, 3.13 months; reading comprehension—handwriting, 5 months; arithmetic—reading comprehension, 7.12 months; spelling—arithmetic, 7.83 months; arithmetic—reading rate, 10.25 months; arithmetic—handwriting, 12.12 months; spelling—reading comprehension, 14.95 months; spelling—reading rate, 18.08 months; spelling—handwriting, 19.95 months. The differences range from 1.87 months (reading rate—handwriting) to 19.95 months (spelling—handwriting). The last four differences show a relative retardation of more than one year and each of them is statistically reliable. The remaining differences are perhaps too small to have much significance and they are, moreover, not entirely statistically reliable.

To summarize the inter-test comparisons on achievement: The differences in achievement measures indicate that twelve-year old rural Negro children are most efficient on spelling, arithmetic, and reading comprehension and least efficient on handwriting and reading rate. From an inspection of the daily programs of the various schools, observing the teachers at work, and watching the subjects during testing periods, the writer found some non-quantitative data which perhaps accounts for the marked differences on the school subjects.

The daily programs have longer drill periods for spelling, arithmetic and reading than for handwriting. In the average Negro school, penmanship is left out of the daily schedule. Arithmetic is a practical subject, which will have eventual value for the Negro. The teacher, realizing the fact, stresses the fundamental processes. Arithmetic has been found to be a popular school subject, therefore, it commands interest and attention. It is small wonder that the pupils have made more advancement in arithmetic than in handwriting. Spelling is a subject that yielded a comparatively high mean score. The differences found here evidently have two sources, namely: interest in subject, as well as drill therein; and also an indirect factor coming through the method of teaching reading. The average Negro school stresses spelling to prepare the child to correspond and read. The "A B C method'' of teaching reading is used to a great extent in the average rural Negro school. The child learns to recognize letters, and not words and phrases. Consequently, the stress placed upon *spelling out* the word to get the meaning makes the child become familiar with ordinary words, and drill by this method causes the child to learn the sound of the word and how it is spelled. This method of reading has in turn negatively influenced the ability to read rapidly. The mean scores show a smaller score in reading rate, yet a fairly high score in reading comprehension. The subjects cannot read as rapidly by this method of spelling out words, but comprehend relatively a larger portion of what they read. This may be due to the relatively small amount read in the given time.

SOME INTER-CORRELATIONS

Treatment of the data so far has been concerned with the amount and reliability of differences in the mean scores on the various measures. In order to show the inter-relationships of the several measures coefficients of correlation (product-moment) were calculated. For a correlation to be considered reliable the obtained **r** should be at least four times its P. E. (Garrett, 1931, p. 170).

It is the investigator's present concern to show: (1) The relationship of intelligence (as measured by Arthur, Stanford, and Illinois) to achievement (as measured by the tests on reading rate, reading comprehension, arithmetic, spelling, and handwriting); (2) the relationship between the separate intelligence measures, and (3) the relationship of each achievement test to each other achievement test. We shall also attempt to account for the relationships found. The several correlations (Table III) are based on ages rather than quotients.

Educational and psychological authorities have different methods of interpreting a given correlation. Perhaps the most authoritative interpretation in the literature is the following: when **r** is less than .15 to .20 it is ''negligible'' or "indifferent;'' when **r** is .15 or .20 to .35 or .40 it is interpreted as a "present but low'' correlation; when **r** is .35 or .40 to .50 or .60 it is "markedly present'' or "marked;'' when **r** is above .60 or .70 it is called a "high'' coefficient of correlation. (Rugg, 1917, p. 256).

Out of the thirty-six correlations given in Table III, twenty-eight are reliable correlations. Six of the eight incompletely reliable correlations involve chronological age. The other two unreliable correlations are: Arthur with reading-rate (.327 ± .095); and, Arthur with handwriting (.224 ± .102). The relation of chronological age to each of the several measures of intelligence follows: Arthur, .457±.084; Stanford, .427±.087, Illinois, .319±.095. Arthur is the most valid measure of the intelligence of our subjects if correlation with chronological age be accepted as a criterion of validity. The relation of chronological age to each of the achievement measures follows: reading rate, .221±.101; reading comprehension, .329±-

TABLE III.

INTER-CORRELATION OF THE SEVERAL MEASURES. TESTS ARE NUMBERED ACROSS THE TOP AS THEY ARE LISTED DOWN THE SIDE OF THE PAGE. THE PROBABLE ERROR IS INDICATED UNDERNEATH EACH COEFFICIENT. N 40.

	2	3	4	5	6	7	8	9	Av. I.	Av. II.	Av. III.
1. Chronological Age	.457* .084	.427* .087	.319 .095	.320 .095	.178 .103	.329 .095	.221 .101	.178 .103	.304	....	....
2. Arthur M. A.	....	.760* .045	.465* .084	.402* .089	.498* .080	.396* .089	.327 .095	.224 .102	.437	.369	....
3. Stanford-Binet M. A.	....	....	.755* .045	.630* .065	.688* .060	.600* .068	.613* .068	.461* .084	.617	.598	....
4. Illinois M. A.	....	....	....	.789* .041	.839* .031	.781* .041	.697* .054	.681* .057	.664	.757	....
5. Spelling E. A.	....	....	....	....	.696* .054	.675* .057	.669* .060	.820* .034	.625	....	.715
6. Arithmetic E. A.	....	....	....	....	....	.531* .077	.486* .082	.721* .051	.580	....	.609
7. Reading Comprehension E. A.	....	....	....	....	....	....	.802* .038	.579* .070	.587	....	.647
8. Reading Rate E. A.	....	....	....	....	....	....	....	.588* .070	.550	....	.636
9. Handwriting E. A.	....	....	....	....	....	....	....	....	.532	....	.677

* Reliable correlation.
Av. I. each measure with each other measure.
Av. II. each intelligence test with each achievement test.
Av. III. each achievement test with each other achievement test.

095; arithmetic, .178±.103; spelling, .320±.095; handwriting, .178 ± .103. Reading comprehension and spelling show the most pronounced conformity with chronological age of any of the measures of achievement.

The degree of relationship between the three measures of intelligence follows: Arthur with Stanford, .760±.045; Arthur with Illinois, .465±.084; Stanford with Illinois, .755±.045. Arthur has more in common with Stanford (**r**=.760) than with Illinois (**r**=.465), while Illinois has more in common with Stanford **r**=.755) than with Arthur (**r**=.465). Of the three measures of intelligence, Stanford is a better measure of the samplings' mental abilities, if conformity with the other two tests be taken as criteria. We have previously given reasons, however, favoring the view that the Arthur test is the most valid measure of the native capacity of our subjects.

The relationship of each achievement measure to the four other achievement measures is found in the last column of Table III. The average correlations thus obtained are as follows: reading rate, .636; reading comprehension, .647; arithmetic, .609; spelling, .715; and handwriting, .677. Or, to state it differently, spelling has more in common with the various school attainments than does any other school subject.

The average and individual coefficients of correlation will be discussed in order to show the degree of relationship between intelligence and achievement. The averages in next to the last column of Table III indicate the relation of each intelligence measure to the average of the achievement measures. The relation of each intelligence test to the average of the achievement measures follows: Arthur, .369; Stanford, .598; Illinois, .757. It is obvious therefore either, (1) that the Illinois test is more nearly a measure of achievement than is either the Arthur or Stanford, or (2) that Illinois is more dependent on formal schooling than are Arthur and Stanford, or (3) that Illinois is a more valid test of the ability of our subjects to master formal school material than is either of the other two tests.

The amount of mental retardation here found on the Illi nois test is not only disproportionately much greater than that found on the Stanford and Arthur tests (both of which are reasonably valid measures of general ability), but is also disproportionately much greater than that reported on comparable Negro subjects by other investigators when presumably valid measures of innate capacity are used. It is also generally agreed that group tests (such as the Illinois) "appear to measure cultural advantages . . . . . . as well as any possible existing innate differences in ability." (Peterson and Lanier, 1929, p. 98). The general ability (including the ability to master school subjects) of the samplings may therefore be presumed to be greater than their scores on the Illinois test indicate. That is, the low rating of the samplings on the Illinois tests is in part due to the fact that success on the Illinois test is dependent on ability to read and follow written directions concerning materials mainly of a formal school nature. Inadequacy in such formal training necessarily results in a low score on the Illinois test.

Arthur compared with the achievement scores indicates a "negligible" correlation between Arthur and handwriting; and a "markedly present" relationship between Arthur and the following: reading comprehension, (.396±.089); arithmetic, (.498±.085); and spelling, (.402±.089). The low individual correlations given in Table III thus further indicate that Arthur is less dependent upon formal training than are the Stanford and Illinois tests.

Stanford correlated with Illinois yields a coefficient of .76, which is to be regarded as indicating a "high" degree of relationship. Stanford in relation to the achievement scores gives a "high" coefficient on every measure with the exception of handwriting (.461±.084). The most plausible inferences to be drawn therefore are: (1) Either that Stanford is a rather highly efficient instrument for measuring ability to master school subjects, or (2) that scores on the Stanford are influenced by academic attainments. Both factors are doubtless operative, although exact quantitative statement cannot be made on the point from our data. We have previously cited evidence (cf. p. 38) favorable to the view that

Stanford I. Q.'s reflect rather strongly the cultural and academic status of the subject.

Illinois gives "high" coefficients of correlation with each of the achievement measures; handwriting (.681) and reading rate (.697) are found to be the lowest but even they are comparatively high. These individual correlations further support the conclusion that the Illinois test is more dependent upon school attainment than is either the Arthur or Stanford tests. The reasons are clear and simple; first of all, Illinois depends to a great degree upon school training, since the ability to follow written directions having to do mainly with verbal material is assumed. Success on an examination of that nature would consequently make the scores higher for each of the achievement measures. In this case failure on Illinois would necessarily lead one to expect a low rank on the achievement measures.

From the data at hand it is concluded that the Arthur scale is a reasonably valid measure of the mental ability (non-verbal) of twelve-year old rural Negro children (**r**=.457), and is but little dependent upon achievement in school. Stanford is a less valid measure of native capacity (**r**=.427) than is Arthur for our subjects, being more dependent upon formal schooling and familiarity with white culture. The Illinois is held to be a less valid measure of native capacity (**r**=.319) of our subjects than is either Arthur or Stanford for reasons previously mentioned. These conclusions are based on a knowledge of the nature and type of each of these measures of general intelligence and upon a consideration of the educational and cultural status of our subjects. The rural Negro, under the conditions of this study, shows much more aptitude for the concrete, non-verbal type of performance than he does for verbal or abstract activities, as measured by the Stanford or Illinois tests. The latter two measures depend to a great extent upon school training, whereas the performance scale does so very slightly, if at all. It is probable that the school is partly responsible for the low rating on both the Stanford and Illinois tests. The results obtained may also be due in part to actual differences in the type of ability (verbal vs. non-verbal) measured by these tests.

That statistically reliable differences in such abilities are found under the conditions of this study is not to be taken as proof of innate differences in verbal and non-verbal ability among Negroes. It may well be that Negro culture, as it now exists, places greater premium on non-verbal activities, (thus leading to greater specialization in such performances to the neglect of more strictly verbal responses) to account for the differences found, rather than that innate differences in the capacities actually exist.

IX. SUMMARY AND CONCLUSIONS

The present study was made for the purpose of comparing the intelligence and achievement of twelve-year old rural Negro children in the schools of Clarke County, Georgia.

To this end forty twelve-year old rural Negroes were selected from the six "best" (out of thirteen) rural schools for Negroes in Clarke County. The subjects were given two individual tests of general intelligence (Arthur and Stanford); one group test of intelligence (Illinois); and, four group achievement tests (Monroe's Reading and Arithmetic, and Ayer's Spelling and Handwriting).

The three tables presented in this study appear to warrant the following conclusions:

(1) Mean score differences indicate that the subjects under study score higher on Arthur, (mean score, 125.25 months) than on either the Stanford, (mean score, 114.50 months), or the Illinois, (mean score, 100.13 months).

(2) Mean scores (stated in months) on the school achievement scales show the following order of proficiency of our subjects in educational attainment: Spelling, (107.33); arithmetic, (99.50); reading comprehension, (92.38); reading rate, (89.25); handwriting, (87.38). The average chronological age is 149.95 months.

(3) Analysis of the amount and reliability of the differences between the separate measures shows:

(a) Using chronological age as a criterion of expected mental ability, our subjects are reliably retarded in the following amounts: as measured by Arthur, 24.45 months, as measured by Stanford, 35.45 months, and as measured by Illinois, 49.72 months. Completely reliable differences are found between chronological age and each of the achievement ages, showing retardation ranging from three and one-half years (spelling) to more than five years (handwriting).

(b) Using the Arthur mental age as a criterion of potential ability to master school subjects we find the amount of educational retardation ranges from more than one and one half years on spelling to more than three years on handwriting. The differences are not only large but are completely reliable. Using Stanford mental age as a criterion the educational retardation is eleven months less than that indicated by Arthur. These differences, although less than Arthur, are completely reliable. Using Illinois mental age as a criterion the education retardation is 25.37 months less than the retardation measured by the Arthur mental age and 14.37 months less than the retardation measured by Stanford mental age In this comparison retardation was shown in all school subjects except spelling, but only two of the differences are completely reliable. A general comparison of this nature seems to indicate that, for our subjects, the Arthur test is the more nearly accurate measure for their potential abilities.

(c) If the relative attainments in subject matter of twelve-year old rural Negro children be compared, we find that our subjects show marked irregularity in their school attainments, being significantly and reliability most proficient in spelling, (mean score, 107.33 months), and least proficient in handwriting, (mean score, 87.38 months).

(4) Inter-correlations of the several measures show that if conformity with chronological age be accepted as a criterion, the Arthur (**r**=.46) and Stanford (**r**=.43) tests are fairly valid measures of general intelligence, and that the Illinois test (**r**=.32) is not as valid a measure of general intelligence as are the Arthur and Stanford tests. If the relation of the intelligence tests to each other be noted, it is found that Arthur has much more in common with Stanford (**r**=.760) than with

Illinois (r=.465) and that Illinois has much more in common with Stanford (r=.755) than with Arthur (r=.465). Stanford, however, has practically the same relation to Illinois (r=.755) as to Arthur (r=.760). Of the three intelligence tests, Illinois is clearly more dependent upon a mastery of school subjects (see average II in the next to the last column of Table III) than is either Arthur or Stanford, while the Arthur is the least dependent upon formal training of a school nature.

When the relations of the achievement tests to one another are studied it is noted that spelling has more in common with each of the measures than does any other subject, while of all the measures, arithmetic has least in common with the other subjects. This is, of course, to be expected from the nature and selection of the tests given.

Following the procedure set out in the statement of the purpose and plan of the study the investigator has found, by the administration of various measures of intelligence and school achievement, that the twelve-year old rural Negro child of Clarke County, Georgia, is mentally capable of a considerably higher degree of school success than has been attained, but because of conditions within the home and within the school he has not had sufficient opportunity or motivation to utilize this mentality in the direction of school attainment. Apparently he shows more aptitude under existing conditions (whether from the influence of environment or heredity or both cannot from our data be definitely and exactly stated) for non-verbal activities than for activities involving abstract, conceptual thinking.

It is hoped that this study may bring to the attention of persons interested in racial problems a quantitative indication of what enlightened public opinion has arrived at, namely: That the Negro child is capable of a reasonably high degree of school success as well as a reasonably higher degree of adaptation to new situations.

Before any widespread practical applications are made of these findings they should be supplemented by further and more extensive research somewhat along the line of this study. It is possible, however, that even the small amount of data here given have sufficient statistical reliability to warrant certain changes of procedure in rural Negro schools. Dr. Grace Arthur, Psychologist to the Child Guidance Clinic of St. Paul, Minnesota, has critically examined our findings and has been kind enough to suggest what she considers to be practical applications of the data. Among other things she suggests the following:* (1) A carefully organized pre-primer course; (2) the utilization of concrete material and of motor activity in all lines of teaching; (3) the maintaining of grade standards by the annual application of an achievement scale; (4) the diverting of those who have reached their limit of achievement along academic lines into the field of occupational training in order to: (a) Relieve academic classes from the burden of an unsuccessful struggle, and (b) to train whatever non-verbal abilities the individual may have. Dr. Arthur points out that this may be far too elaborate a program to be practicable at the present moment, but that it might well be held in mind as an objective toward which to work. With each of these views of Dr. Arthur the writer finds herself in agreement.

---

* Personal letter from Dr. Arthur.

BIBLIOGRAPHY

1. Arlitt, Ada, "Relation of Intelligence to Age in Negro Children." *Journ. Appl. Psychol.*, 1921, 5. 378-384.
2. Arthur, Grace, *A Point Scale of Performance*, New York, The Commonwealth Fund Division of Publications, 1930.
3. Bigge, L. A. Selby, *Report on Psychological Tests*, English Edition, London, His Majesty's Stationery Office, 1924.
4. Binet, A., "Les Idees Sur Les Enfants." 1909.
5. Binet, A., and Simon, Th., "L'intelligence des imbeciles," *L'Annee Psychologique*, 1909, pp. 1-147.
6. Duggan, M. L., *Fifty-Eighth and Fifty-Ninth Annual School Report*, Atlanta, Georgia, 1930.
7. Freeman, F. N., Holzinger, K. J., Mitchell, B. C., "The Influence of Environment on the Intelligence, School Achievement, and Conduct of Foster Children." *Yearbook for the Study of Education*, Vol. 27, 103, 217. 1928.
8. Garrett, H. E., *Statistics in Psychology and Education*, New York, Longman's, Green and Company. 1931.
9. Holzinger, K. J., *Statistical Tables for Students in Education and Psychology*. University of Chicago Press. 1925.
10. Isserlis, L., "The Relation Between Home Conditions and the Intelligence of School Children," *Great Britain Medical Research Council, Special Report Series*. No. 74, 1923, 17-18.
11. Kelly, T. L., *Interpretation of Educational Measurements*, Yonkers, World Book Company. 1927.
12. Levine, A. J., and Marks, L., *Testing Intelligence and Achievement*. New York, The Macmillan Company, 1928.
13. Monroe, M., *Diagnostic Reading Examination*. Chicago Institute of Juvenile Research. 1930.
14. Monroe, W. S., and Buckingham, B. R., *The Illinois Examination I and II Teacher's Handbook*. The Public School Publishing Company, Bloomington, Ill. 1921.
15. Otis A. S., "Comparison of Starch and Ayers Tests." *School and Society*, 1916, 4. 753-756.
16. Peterson, J., Lanier, L. H., and Walker, H. M., "Comparison of Whites and Negroes in Certain Ingenuity and Speed Tests," *Journ. Comp. Psychol*, 1925, 3, 271-283.
17. Peterson, J., Lanier, L. H., "Studies in the Comparative Abilities of Whites and Negroes." *Mental Meas. Monog.*, 1929, No. 5.
18. Pyle, W. H., "The Mind of the Negro Child." *School and Society*. 1915, I. 357-360.
19. Ruch, G. M., and Stoddard, G. D., *Tests and Measurements in High School Instruction*. Yonkers, World Book Company. 1927.

20. Rugg, H. O., *Statistical Methods Applied to Education.* Boston, Houghton Mifflin, 1927.

21. Schluter, W. C., *How to do Research Work*, New York. Prentice-Hall, Inc., 1926.

22. Stoelting, C. H., Co., Psychological and Physiological Apparatus and Supplies, Chicago, 1932.

23. Terman, L. M., *The Measurement of Intelligence*, Boston, Houghton Mifflin, 1916.

24. Thomas, W. I., and Thomas, D. S., *The Child in America, New* York, Alfred A. Knopf, 1928.

APPENDIX A.

RAW SCORE DATA ON EACH PUPIL ON THE VARIOUS MEASURES EXPRESSED IN MONTHS AND IN EDUCATIONAL AND INTELLIGENCE QUOTIENTS AND MENTAL AGES. N 40 CASES.

1*	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25
1....	148	131	116	114	90	102	102	109	86	89	78	80	61	78	70	90	69	86	75	96	58	74	3	F
2....	156	152	140	126	96	108	138	133	117	97	90	81	62	69	69	77	88	98	85	94	75	83	5	M
3....	153	148	107	90	78	84	90	84	74	97	70	61	51	73	55	79	59	84	55	79	48	69	2	F
4....	145	90	89	78	84	90	84	79	72	60	61	55	58	95	63	104	58	95	54	89	50	82	1	F
5....	156	118	111	96	84	84	90	82	75	77	71	62	54	76	54	76	58	82	53	76	48	68	1	F
6....	149	149	139	144	96	102	156	135	123	100	93	100	64	69	68	74	104	112	91	100	83	89	5	F
7....	145	156	122	102	84	90	90	96	74	108	84	71	58	69	63	75	63	75	66	79	51	61	2	M
8....	145	133	116	96	84	84	102	96	77	92	80	67	58	73	58	73	70	88	66	83	53	66	2	M
9....	150	124	112	90	96	96	96	83	74	83	75	61	64	85	64	85	64	85	55	73	49	65	2	M
10....	155	95	90	72	78	84	84	84	72	63	58	49	50	86	54	93	54	93	54	93	46	79	2	F
11....	144	98	104	72	72	84	84	94	78	68	72	51	50	69	58	81	58	81	65	90	54	75	2	F
12....	144	106	114	78	78	84	84	86	81	74	80	55	54	68	58	73	58	73	60	75	56	70	4	F
13....	144	90	96	72	72	78	84	79	72	63	67	51	50	75	54	81	58	87	55	81	50	75	2	M
14....	145	94	90	84	78	84	90	96	81	65	62	59	54	87	58	94	63	102	66	106	56	90	3	F
15....	146	119	106	90	84	84	96	86	85	82	73	63	58	79	58	79	66	90	59	81	58	79	3	M
16....	146	109	98	90	84	84	90	100	78	75	67	63	58	87	58	87	62	93	68	101	53	79	3	F
17....	154	104	94	90	78	84	96	101	77	68	68	61	51	75	55	81	52	76	66	97	50	74	4	M
18....	145	125	92	84	84	84	90	100	85	86	63	59	58	92	58	92	63	100	69	110	59	94	3	M
19....	145	122	102	96	84	84	96	114	85	84	70	67	58	83	58	83	66	94	79	113	59	84	4	F
20....	155	120	116	84	84	96	90	110	78	77	75	57	54	72	62	83	59	78	71	95	50	67	4	M
21....	150	138	137	144	96	108	132	138	100	92	91	96	64	70	72	79	96	105	92	101	67	74	7	F
22....	156	146	131	126	108	120	108	131	105	94	84	81	69	82	77	92	69	82	84	100	67	80	5	F

APPENDIX A—(Continued)

23....152	142	126	96	90	90	96	113	83	93	83	65	59	71	59	71	63	76	74	89	55	66	3	F
24....156	178	142	120	90	90	102	118	84	114	91	77	58	64	58	64	65	71	77	85	54	59	4	F
25....154	141	125	96	84	96	90	109	78	92	81	65	55	68	52	64	58	72	71	88	51	63	3	M
26....150	122	109	108	84	84	102	119	100	81	73	73	56	77	56	77	68	93	79	108	67	92	5	M
27....150	142	112	102	84	90	102	108	79	95	75	69	56	75	60	80	68	91	72	96	53	71	3	M
28....145	119	119	96	96	96	102	86	79	82	82	67	66	80	66	80	70	85	59	72	54	66	3	M
29....155	133	114	96	90	96	96	109	79	86	74	65	59	80	62	84	62	84	70	95	51	69	3	M
30....155	133	142	132	108	108	126	127	100	86	92	88	70	76	70	76	81	88	82	89	65	71	6	F
31....144	127	114	108	90	96	102	114	110	88	79	76	67	85	67	85	71	90	79	100	76	84	3	F
42....149	138	112	96	90	90	90	109	91	93	75	67	60	80	60	80	60	80	73	97	61	81	3	F
33....146	132	111	78	78	84	90	94	73	90	76	55	53	70	58	76	62	82	64	84	50	66	3	M
34....149	147	125	132	120	132	108	134	100	99	84	92	81	96	89	106	72	86	90	107	67	80	5	F
35....146	126	117	120	102	90	96	112	84	86	80	84	70	88	62	78	66	83	77	96	58	73	3	F
36....152	122	108	114	90	90	102	127	121	80	71	76	59	83	59	83	67	94	84	118	80	113	5	F
37....146	104	116	96	114	114	96	121	108	71	79	67	78	99	78	99	66	84	83	105	74	94	5	F
38....148	105	117	84	102	84	84	98	77	71	79	59	70	89	57	72	57	72	66	84	52	66	3	M
39....150	128	119	84	84	90	90	85	74	85	79	57	56	71	60	76	60	76	57	72	49	62	3	M
40....155	103	110	108	96	108	96	129	106	66	71	73	62	87	70	99	62	87	83	117	68	96	4	F

* The several raw scores and additional data indicated by numbers as follows:

1. Pupil's number.
2. Chronological age
3. Arthur Pt. Perf. M. A.
4. Stanford-Binet M. A.
5. Illinois Gen. M. A.
6. Ill. Read. Rate E. A.
7. Ill. Read. Comp. E. A.
8. Ill. Arithmetic E. A.
9. Ayer's Spelling E. A.
10. Ayer's Handwriting E. A.
11. Arthur Pt. Perf. I. Q.
12. Stanford-Binet I. Q.
13. Ill. Gen. I. Q.
14. Ill. Read. Rate E. Q.
15. Ill. Read. Rate A. Q.
16. Ill. Read. Comp. E. Q.
17. Ill. Read. Comp. A. Q.
18. Ill. Arithmetic E. Q.
19. Ill. Arithmetic A. Q.
20. Ayer's Spelling E. Q.
22. Ayer's Handwriting E. Q.
23. Ayer's Handwriting A. Q.
24. Grade.
25. Sex.

APPENDIX B.

TABLE SHOWING THE EXTENT TO WHICH THE SUBJECTS ARE UNDER-PLACED, CORRECTLY-PLACED, OR OVER-PLACED AS ESTIMATED BY THE VARIOUS CRITERIA OF PLACEMENT. N 40.

Criteria of Placement	Grade I N=2	Grade II N=7	Grade III N=16	Grade IV N=6	Grade V N=7	Grade VI N=1	Grade VII N=1
Arthur M. A.	Under-placed	Under-placed	Under-placed	Correctly placed	Correctly placed	Over-placed	Over-placed
Stanford M. A.	Under-placed	Under-placed	Correctly placed	Over-placed	Over-placed	Over-placed	Over-placed
Illinois M. A.	Under-placed	Correctly placed	Over-placed	Over-placed	Over-placed	Over-placed	Over-placed
Average M. A.	Under-placed	Under-placed	Correctly placed	Over-placed	Over-placed	Over-placed	Over-placed
Reading Rate E. A.	Under-placed	Correctly placed	Over-placed	Over-placed	Over-placed	Over-placed	Over-placed
Reading Comprehension E. A.	Under-placed	Correctly placed	Over-placed	Over-placed	Over-placed	Over-placed	Over-placed
Arithmetic E. A.	Under-placed	Correctly placed	Over-placed	Over-placed	Over-placed	Over-placed	Over-placed
Spelling E. A.	Correctly placed	Correctly placed	Over-placed	Over-placed	Over-placed	Over-placed	Over-placed
Handwriting E. A.	Correctly placed	Correctly placed	Over-placed	Over-placed	Over-placed	Over-placed	Over-placed
Average E. A.	Under-placed	Correctly placed	Over-placed	Over-placed	Over-placed	Over-placed	Over-placed

PHELPS-STOKES FELLOWSHIP STUDIES

1. **Woofter, Thomas J. Jr.** *Negroes of Athens, Georgia.* 1913. 62 p. illus. map. 40 cents.
2. **Hill, Walter B.** *Rural Survey of Clarke County, Georgia, with Special Reference to the Negroes.* 1915. 63 p. illus. maps. 40 cents.
3. **Johnson, Miley K.** *School Conditions in Clarke County, Georgia, with Special Reference to the Negroes.* 1916. 50 p. illus. maps. 35 cents.
4. **O'Kelley, Hoke S.** *Sanitary Conditions Among the Negroes of Athens, Georgia.* 1918. 30 p. 25 cents.
5. **Long, Francis T.** *Negroes in Clarke County, Georgia, During the Great War.* 1919. 56 p. 35 cents.
6. **Reed, Ruth.** *The Negro Women of Gainesville, Georgia.* 1921. 61 p. illus. 40 cents.
7. **Scarbrough, Dorothy D.** *An Economic Study of Negro Farmers as Owners, Tenants and Croppers.* 1924. 37 p. illus. 25 cents.
8. **Patrick, James R.** *A Study of Ideals, Intelligence and Achievements of Negroes and Whites.* 1926. 48 p. illus. diagrams. 40 cents.
9. **Fanning, John W.** *Negro Migration.* 1930. 39 p. diagr. maps. 40 cents.
10. **Wilson, Ernest B.** *The Water Supply of the Negro.* 1930. 80 p. illus. 80 cents.
11. **Wardlaw, Ralph.** *Negro Suffrage in Georgia, 1867-1930.* 1932. 91 p. 60 cents.
12. **Thurmond, Sarah.** *A Comparison of the Intelligence and Achievement of Twelve-Year Old Negro Children in the Rural Schools of Clarke County, Georgia.* 1933. 44 p. 50 cents.